BNL's Role in US ATLAS Computing

Torre Wenaus, BNL

DOE Annual HEP Program Review
BNL
April 18, 2006
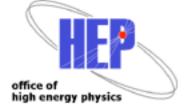
ATLAS Computing Timeline

1999	**Physics TDR** based on previous generation software (Geant3, Fortran)
2000+	**New C++ software:** Athena framework, Geant4, reconstruction, …
2002	**LHC Computing Grid (LCG) launched** to provide facilities, software
2003	**Data Challenge 1:** First large scale use of the grid; Geant3 based
2004	**Data Challenge 2:** Fully grid based; Geant4, LCG software
2004	**Combined test beam:** Real data analysis with new software, ongoing
2005	**Rome physics workshop:** First large scale analysis
2005	**ATLAS computing TDR** laying out the computing model
2006	**Computing System & Detector Commissioning**
Fall 2006	**LCG computing infrastructure for startup complete**
Fall 2006	**Cosmic ray run**
Spring 2007	**Datataking rehearsal** (proposed)
Late 2007	**Datataking**

ATLAS Data Flow 2008

From online: 2.7 TB/day
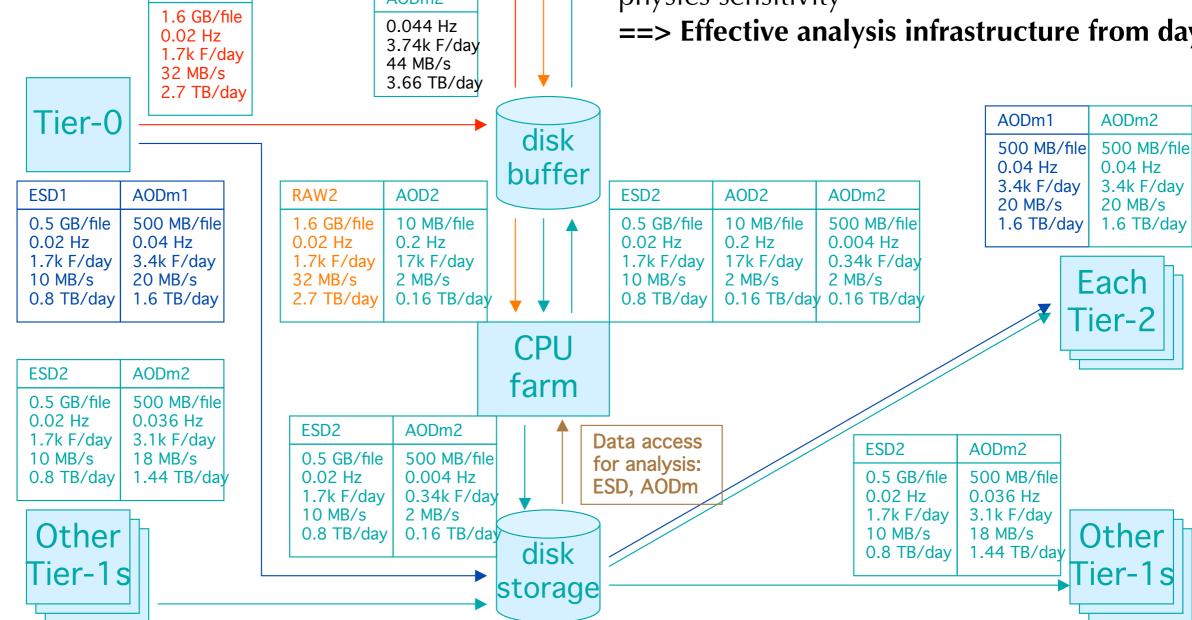To CERN tape: 3.7 TB/day
Tier 0 disk to BNL tape: 200 MB/s

During early running:
Low luminosity (~10% nominal) but
loose triggering for performance studies
==> Nominal data volumes from day 1
Time critical detector validation and early
physics sensitivity
==> Effective analysis infrastructure from day 1

Tape

RAW1
ESD2
AODm2
0.044 Hz
3.74k F/day
44 MB/s
3.66 TB/day

RAW1
1.6 GB/file
0.02 Hz
1.7k F/day
32 MB/s
2.7 TB/day

Tier-0

disk buffer

ESD1	AODm1
0.5 GB/file	500 MB/file
0.02 Hz	0.04 Hz
1.7k F/day	3.4k F/day
10 MB/s	20 MB/s
0.8 TB/day	1.6 TB/day

RAW2	AOD2
1.6 GB/file	10 MB/file
0.02 Hz	0.2 Hz
1.7k F/day	17k F/day
32 MB/s	2 MB/s
2.7 TB/day	0.16 TB/day

ESD2	AOD2	AODm2
0.5 GB/file	10 MB/file	500 MB/file
0.02 Hz	0.2 Hz	0.004 Hz
1.7k F/day	17k F/day	0.34k F/day
10 MB/s	2 MB/s	2 MB/s
0.8 TB/day	0.16 TB/day	0.16 TB/day

AODm1	AODm2
500 MB/file	500 MB/file
0.04 Hz	0.04 Hz
3.4k F/day	3.4k F/day
20 MB/s	20 MB/s
1.6 TB/day	1.6 TB/day

Each Tier-2

ESD2	AODm2
0.5 GB/file	500 MB/file
0.02 Hz	0.036 Hz
1.7k F/day	3.1k F/day
10 MB/s	18 MB/s
0.8 TB/day	1.44 TB/day

CPU farm

ESD2	AODm2
0.5 GB/file	500 MB/file
0.02 Hz	0.004 Hz
1.7k F/day	0.34k F/day
10 MB/s	2 MB/s
0.8 TB/day	0.16 TB/day

Data access
for analysis:
ESD, AODm

ESD2	AODm2
0.5 GB/file	500 MB/file
0.02 Hz	0.036 Hz
1.7k F/day	3.1k F/day
10 MB/s	18 MB/s
0.8 TB/day	1.44 TB/day

Other Tier-1s

disk storage

Other Tier-1s

ATLAS Computing at BNL

- Principal objectives:
  - Fulfill role as principal US center (Tier 1) in the tiered ATLAS (and LHC) computing model
    - Supply capacity to ATLAS as agreed in MOU (~20%)
    - Guarantee the capability and capacity needed by the US ATLAS physics program
  - Establish a critical mass in computing and software to support ATLAS physics analysis at BNL and elsewhere in US
  - Contribute to ATLAS software most critical to enabling physics analysis at BNL and the US
  - Help US ATLAS physicists to avail themselves of BNL expertise and facilities and strengthen US ATLAS physics
  - Leverage projects outside ATLAS, eg the grid projects, where they can strengthen the BNL and US ATLAS programs

ATLAS Computing at BNL

- Principal activities:
  - US ATLAS Tier 1 Center
    - Tier 1 facility development and operations
    - Tier 2 support: services, expertise, coordination, workshops
    - Grid computing development, collaboration
  - ATLAS core software
    - Software support and QA
    - Data management
    - Event data model and event store
    - Distributed production systems
    - Analysis tools and distributed analysis
  - ATLAS detector and physics analysis software (cf. Kyle's talk)
    - Calorimeter, muons, trigger subsystems
    - Combined reconstruction
    - Higgs, SUSY and Heavy Ion analysis participation

BNL in ATLAS Computing Leadership

BNL holds leadership roles in all principal activity areas

- K. Assamagan, Physics Analysis Tools Coordinator
- B. Gibbard/R. Popescu, US ATLAS Facilities Managers
- A. Klimentov, Software Integration Group Coordinator
- A. Klimentov, DDM Operations Coordinator
- S. Rajagopalan, Calo software/performance Co-Coordinator
- S. Rajagopalan, US ATLAS Software Manager
- A. Undrus, US ATLAS Software Support Manager
- T. Wenaus, Data Management/Database Project Co-Leader
- T. Wenaus, Panda Project Co-Leader
- T. Wenaus, US ATLAS Distributed Software Manager

US in the ATLAS Computing Model

- BNL Tier 1 responsibilities (~10 Tier 1s in all)
  - Archival shares of raw and reconstructed data, and associated calibration & reprocessing
  - Store and serve 100% of ATLAS reconstruction (ESD), analysis (AOD) and physics tag data (TAG)
  - Physics group level managed production/analysis
  - Resources dedicated to US physicists: additional per-physicist capacity at 50% of the level managed centrally by ATLAS)
- US Tier 2s have complementary role
  - Bulk of simulation and end user analysis support
  - Store and serve 100% of AOD, TAG and subset of ESD
- Tier 1 and Tier 2s both support institutional and individual users
  - Primarily end user analysis

BNL Tier 1 Facility Evolution

- 2005
  - ATLAS staff increased 4 FTEs to 11.5 at year end
    - Ramp-up in support of databases, data storage/distribution, grid administration/operations, distributed production
  - 4-fold increase in CPU to 740 kSi2k
  - dCache in production, managing ~200TB of farm disk
- 2006
  - Staff increase to 15 FTEs (data storage & management, fabric & grid infrastructure, operations & user support)
  - US operations hub for new data management and production/analysis systems
    - ~50% of US CPUs currently in production are at BNL
  - HPSS mass storage upgrade (just completed)
  - WAN/LAN upgrade to 10 Gb/s (done)
  - 200 MB/s required for CERN disk - BNL tape *exceeded* (this month)
  - CPU to ~1200 kSi2k
  - Disk to ~500 TB

US Tier 1 and Tier 2s

BNL Tier 1

2006: 1200 kSi2k CPU, 500 TB disk
2008: 8000 kSi2k CPU, 4600 TB disk, 8 PB tape

Northeast Tier 2
(BU, Harvard)

2006: 350 kSi2k CPU, 170 TB disk
2008: 1090 kSi2k CPU, 480 TB disk

Midwest Tier 2
(U Chicago, Indiana)

2006: 510 kSi2k CPU, 130 TB disk
2008: 1100 kSi2k CPU, 465 TB disk

Southwest Tier 2
(UT Arlington, Oklahoma U, UNM, Langston)

2006:900 kSi2k CPU, 200 TB disk
2008: 1700 kSi2k CPU, 540 TB disk

Two more Tier 2 selections July 2006

2008 Tier 2 totals:
6500 kSi2k CPU, 2900 TB disk

In 2008 ~60% of CPU & disk and 100% of tape storage will be at BNL

office of
high energy physics

ATLAS Software at BNL

- A collaboration of three groups in the Physics Department
  - **Physics Applications Software:** D. Adams, W. Deng, A. Klimentov, T. Maeno, P. Nevski, M. Nowak, A. Undrus, T. Wenaus
    - Core software, user support, production operations
  - **Omega:** K. Assamagan, K. Cranmer, H. Ma, S. Rajagopalan, S. Snyder et al
    - Physics analysis software and tools, core software, user support, analysis
  - **Tier 1 Facility:** J. Hover, Y. Smirnov, X. Zhao et al
    - DB services, facility integration, core software
    - Close interactions between software and facility people

Software Support and Quality Assurance

- Alex Undrus is US ATLAS Librarian, US responsible for software support, and is a principal contributor to ATLAS software infrastructure and QA support
  - ATLAS software installations at BNL Tier 1 for US use
  - ATLAS-wide automated nightly software builds and QA test suites
  - US ATLAS software user support
  - Hire is pending to expand support for ATLAS software users and Tier 2 software deployment
- David Adams is responsible for data validation in US ATLAS production
- Pavel Nevski (@CERN) is responsible ATLAS-wide for production task definition and associated QA

Data Management

- BNL leads the effort at the ATLAS and US levels
  - Torre Wenaus co-leads the ATLAS Database and Data Management Project and leads the Distributed Data Management subproject
  - Alexei Klimentov (@CERN) leads Distributed Data Management Operations for ATLAS
    - Deploying the new DDM to all ATLAS Tier 1s
  - Wensheng Deng is responsible for Distributed Data Management Operations for US ATLAS
  - We work closely with the core team in the CERN ATLAS group developing the system, Don Quixote 2 (DQ2)
- The last year has been extremely productive: a fully re-engineered system developed, deployed, validated, and now in production (BNL and the US first in deployment)

Event Data Model and Event Store

- BNL has been a leader from the beginning on the ATLAS event data model, its in-memory representation and its storage
- Srini Rajagopalan, Hong Ma originated the ATLAS transient event model (StoreGate) and continue as experts/developers
- In the last ~year new BNLers have made major contributions:
  - Scott Snyder (D0 event model expert) has become a major contributor to event model development
  - Kyle Cranmer has led development of the 'event view' event data representation for analysis and wrote most AOD classes
  - Tadashi Maeno has led development of event data representations for final analysis
  - Marcin Nowak (@CERN) now the ATLAS expert on the internals of the ATLAS event store (ROOT I/O and its use in POOL)
  - Kyle and Marcin on Event Management Board
- One pending BNL hire to connect the effort to data management

Distributed Production

- As in DDM, the last year has seen a complete re-engineering of distributed production
    - With BNL stepping into major roles (BNL was not involved in development of the previous generation systems!)
- Panda - Production ANd Distributed Analysis System
    - Initiated as a complete redesign in Aug 2005
    - Co-led by T. Wenaus, lead developer T. Maeno, and several other BNLers on the team (D. Adams, W. Deng, Y. Smirnov, X. Zhao); BNL responsible for ~8 of 10 system components
    - Rapid development led to Panda take-over of US ATLAS production in Dec
    - Fully integrated with ATLAS DDM
    - Runs as integral part of ATLAS production, and in standalone
    - Requires half the operations manpower and already delivers >4x the scalability of the previous system, with anticipated further factor 10-20
- Planned BNL Panda hire went to UTA because of our high overhead

Distributed Analysis

- A very active year, but not all good!
- BNL gave up ATLAS DA coordination mainly due to the failure of the BNL-based DIAL system (D. Adams) to achieve broad take-up
- Following this, BNL and US ATLAS reorganized (under T. Wenaus) and replanned distributed analysis
- DA is now a principal design goal and deliverable for Panda
- DIAL has been used in debugging, benchmarking and validating Panda, but will subsequently be phased out (development is stopped)
- Panda based analysis came online in October (T. Maeno) and will be deployed at a usable scale (sufficient resources to not be overwhelmed by production) in next few weeks
- Panda and Panda-based analysis have captured interest in wider ATLAS
- Panda now on the 'short list' of analysis systems being implemented by ATLAS behind a generic front end (GANGA)
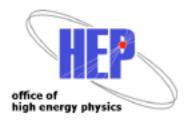
Physics Analysis Tools

- BNL leads the ATLAS Physics Analysis Tools (PAT) effort: Ketevi Assamagan
- BNL also has two of the key ATLAS developers of analysis tools: Kyle Cranmer, Tadashi Maeno
- Benefitting also from the active analysis community at BNL for generating ideas & tools and testing/improving them
- The team has led in the key ATLAS PAT deliverables:
  - Event representation for analysis (AOD)
  - EventView toolkit for using AODs in analysis
  - Athena (Framework) aware Ntuple for ROOT based final analysis
  - Extensive tutorials, documentation

Conclusion

- BNL has continued to focus its computing involvement on areas most critical to BNL and US ATLAS physics analysis
  - Very successfully, by leadership and technical measures
- The BNL Tier 1 serves as the hub and principal center of the US community, with scale-up for datataking underway
- The BNL computing effort provides the critical mass expertise and resources to support BNL as a leading analysis center and US ATLAS as a leader in physics analysis

More Information

- Feb 2006 DOE/NSF review of ATLAS computing
  - http://agenda.cern.ch/fullAgenda.php?ida=a06179
- US ATLAS Physics, Software and Computing
  - http://www.usatlas.bnl.gov/atlas_psc/
- BNL Tier 1
  - http://www.acf.bnl.gov/
- DDM
  - https://uimon.cern.ch/twiki/bin/view/Atlas/DistributedDataManagement
- Panda
  - https://uimon.cern.ch/twiki/bin/view/Atlas/PanDA
- Distributed analysis with Panda
  - https://uimon.cern.ch/twiki/bin/view/Atlas/DAonPanda
- Physics analysis tools
  - https://uimon.cern.ch/twiki/bin/view/Atlas/PhysicsAnalysisTools